

Scheduling

- **What is Scheduling?**
 - Process of assigning Pods to Nodes in a Kubernetes cluster.
- **Kubernetes Scheduler:**
 - The component responsible for deciding which Node a Pod should run on.
 - Works based on the available resources and the Pod's resource requirements.
- **Node Selection:**
 - Factors considered by the scheduler:
 - CPU, memory resources.
 - Node affinity.
 - Taints and tolerations.
 - Pod affinity and anti-affinity.
 - Resource requests and limits.
- **Scheduling Policies:**
 - Default scheduling behavior.
 - Priority scheduling (priority classes).
 - Custom scheduling policies (e.g., using `kube-scheduler` with custom rules).
- **Affinity and Anti-Affinity:**
 - **Node Affinity:** Constraints that allow or prevent Pods from being scheduled on specific nodes.
 - **Pod Affinity:** Ensures that Pods are scheduled on nodes where other specific Pods are running.
 - **Pod Anti-Affinity:** Ensures Pods are not scheduled on the same node as other specified Pods.
- **Taints and Tolerations:**
 - **Taints:** Applied to Nodes to prevent Pods from being scheduled unless they tolerate the taint.
 - **Tolerations:** Allow Pods to be scheduled on Nodes with specific taints.
- **Resource Requests and Limits:**
 - **Requests:** Minimum amount of resources a Pod needs (scheduler uses this to place Pods).
 - **Limits:** Maximum amount of resources a Pod can use (enforced during execution).
- **Preemption:**
 - When a Pod with higher priority displaces lower-priority Pods to be scheduled.
- **Scheduler Extenders:**

- Custom schedulers or extender mechanisms to modify default scheduling decisions.
 - **Pod Disruption Budgets (PDB):**
 - Used to manage voluntary disruptions (like upgrades) to ensure availability during disruptions.
-

Revision #2

Created 18 November 2024 21:51:08 by Admin

Updated 21 November 2024 19:58:58 by Admin